PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/85, 15/63, 15/64, A61K 48/00 — A1

(11) International Publication Number: WO 95/32297

(43) International Publication Date: 30 November 1995 (30.11.95)

(21) International Application Number: PCT/GB95/01195

(22) International Filing Date: 25 May 1995 (25.05.95)

(30) Priority Data:
9410446.0 25 May 1994 (25.05.94) GB

(71) Applicant *(for all designated States except US)*: CANCER RESEARCH CAMPAIGN TECHNOLOGY LIMITED [GB/GB]; Cambridge House, 6-10 Cambridge Terrace, Regent's Park, London NW1 4JL (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: BROWN, William [GB/GB]; 3 The Green, Clanfield, Oxfordshire OX18 2SR (GB).

(74) Agent: GREAVES, Carol, P.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: NOVEL CHROMOSOME FRAGMENT AND ITS USE AS A VECTOR

(57) Abstract

The present invention relates to the field of molecular genetics and in particular to artificial chromosomes, to their preparation in particular using telomere-directed chromosome fragmentation techniques and to their use as DNA vectors, for instance for application in gene therapy and animal transgenesis. Vectors of the invention comprise a chromosome fragment which fragment is at least partly responsible for centromere function of the parent chromosome and is capable of replication and segregation during cell cycle, and which is of such a size that it can be resolved using gel electrophoresis. Suitable fragments are derived from the human Y chromosome.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

Novel chromosome fragment and its use as a vector

The present invention relates to the field of molecular genetics and in particular to artificial chromosomes, to their preparation in particular using telomere-directed chromosome fragmentation techniques and to their use as DNA vectors, for instance for application in gene therapy and animal transgenesis.

Following work with yeast artificial chromosomes (YACs), there is currently widespread interest in the prospect of building artificial chromosome vectors from other hosts such as mammals and plants, for gene therapy and animal transgenesis. In order to achieve useful mammalian artificial chromosomes (MACs), it is necessary to generate chromosomal DNA molecules which preferably retain their linear integrity and which replicate and segregate accurately during cell cycle. It is not presently understood how large a mammalian chromosome has to be in order to achieve this, nor what DNA sequences are necessary to retain full centromere function, although the sequence requirements for mammalian telomere function are understood in outline (Barnett et al., 1993).

In particular it is not generally understood how centromeric DNA interacts with proteins to bring about accurate chromosome segregation nor why, when much of the cell cycle is conserved (Nurse, 1990; Murray and Hunt, 1993) it varies so widely during evolution.

For instance, satellite DNA, composed of tandemly repeating sequences, is found at or close to the centromeres of many multicellular eukaryotes but has often been dismissed as being without functional significance because it is polymorphic and evolves rapidly. However human alphoid satellite DNA

(Manuelidis, 1978) has been directly implicated in centromere function by two types of observation. Alphoid DNA has been introduced into either monkey (Haaf et al., 1992) or hamster chromosomes (Heartlein and Latt, 1988) and appeared to cause the chromosomes to behave as though they were dicentric: they lagged at anaphase and were structurally unstable. The interpretation of these experiments has been controversial as the sequence organization of the integrated alphoid DNA was not described in detail, few independent integration events were analyzed and other non-centromeric sequences (Cooper and Tyler-Smith, 1992) have been reported to show similar properties (Hadlaczky et al., 1991). A second approach to identifying human centromeric DNA has been to map centromeric sequences using re-arranged natural chromosomes; in this way Tyler-Smith and colleagues (1993) mapped the centromeric DNA of the human Y chromosome to a region which included the alphoid DNA and 300kb. of short arm flanking sequence. The observation that centromere function could be suppressed in a "dicentric" Y chromosome by partial deletion of alphoid DNA (Tyler-Smith et al., 1993) implicated alphoid DNA directly, although its functional significance remained undefined.

Insertional mutagenesis using existing vector systems poses problems, particularly where megabase sized fragments of DNA are to be introduced in a defined sequence environment into the genomes of humans and animals.

Use of a vector based upon a stable autonomously replicating chromosome fragment would resolve many of these problems. However it would be expected that fragments of a size which would ensure mitotic stability would too large to be isolated and manipulated in a manner required for a use in current recombinant DNA in-

vitro technology as a basis for a vector.

The present invention provides a vector comprising a chromosome fragment which fragment comprises sequences which have centromeric activity, is capable of replication and segregation during cell cycle, and is of such a size that it can be resolved using gel electrophoresis.

The present invention also provides a method for obtaining a chromosome fragment for use as a vector as described above which method comprises dissecting a chromosome in such a manner that a fragment retaining centromere function is retained.

Methods used in the dissection will vary depending upon the particular chromosome used and the state of knowledge regarding the structure of that chromosome and the location of the centromere region. It may be possible to specifically design dissection techniques which target sites one either side of the centromere region assuming the location and nature of this region is specifically understood. However a particularly useful technique would be the application of telomere directed truncation methods for example as described by Surosky and Tye, 1985 and Barnett et al., 1993. Such techniques involve the use of telomeric DNA to effect a break in the chromosome. The technique may be repeated for as often as required in order to achieve a fragment of the desired size. After each step, the ability of the fragments obtained to replicate and segregate during cell cycle may be confirmed by mitosis experiments and only those which are mitotically stable to a reasonable extent selected. Techniques such as *in situ* hybridisation to chromosomes in whole cells such as described in Funabiki et al. (1993) may be employed in these experiments.

the chromosome, selecting cells which have stably integrated vector, and isolating a clone containing a chromosome derivative which has lost one arm of the chromosome but retained the other arm truncated by telomere-directed breakage within the centromeric array, such clone segregating accurately at mitosis and being retained by cells proliferating in the absence of selection.

The selected clone is then subjected to a further round of telomere-directed breakage to obtain a mini-chromosome fragment which is maintained stably by cells proliferating in culture. The second round breakage construct will suitably comprise telomeric DNA and a marker gene under the control of appropriate control agents such as promoters. An example of a suitable second round breakage construct is a linearised plasmid containing a 1.2kb stretch of human telomeric DNA, an *Escherichia coli* guanine phosphoribosyl transferase gene driven by an SV40 early region promoter and a *Saccharomyces cerevisiae* ADE2 gene.

We have found that in this manner one can obtain a mini-chromosome fragment small enough to be resolved by gel electrophoresis.

The applicants have identified a number of useful chromosome fragments which are derived from the human Y chromosome. The fragments are derived either from the short arm of the chromosome or the long arm of the human

Y chromosome. Preferred small fragments have been isolated from the long arm of the human Y chromosome.

Therefore in a particular embodiment, the present invention provides a vector comprising a linear derivative of the human Y chromosome which includes at least some alphoid DNA, said fragment being capable of replication and segregation during cell cycle, and capable of resolution by gel electrophoresis.

A fragment derived solely from the short arm of the human Y chromosome is approximately 8Mb in size. This fragment is suitably prepared by effecting the steps of carrying out a first telomere-directed truncation of the human Y chromosome, isolating stably segregating a fragment which contains only short arm and alphoid DNA, carrying out a second telomere-directed truncation of the isolated fragment, and isolating a shorter stably segregating fragment. A cell line containing this fragment designated Δ1 was deposited at the European Collection of Animal Cell Cultures (ECACC) on 23 May 1995 under accession number 95052342.

Fragments derived using a similar telomere directed truncation technique have been obtained from the long arm of the human Y chromosome. Fragments of approximately 6Mb in size and also as small as 3.5Mb in size are obtained in this way. The structure of these fragments has not been fully elucidated and it is believed that some form of rearrangement, possibly involving movement of the centomere takes place. However the fragments so obtained are capable of replication and segregation and their size makes them very useful as vectors. Cell lines containing three such fragments designated Δ7, Δ128 and Δ196 were deposited at the ECACC on 23 May 1995 under the accession numbers 95052343, 95052344 and 95052345 respectively.

Δ7 and Δ128 were found to be similar but not identical in size and are approximately 6Mb long. Δ196 was found to be approximately 9Mb long.

In yet a further aspect the invention comprises a vector as described above into which a gene encoding a desired protein has been incorporated. Furthermore the invention provides a cell transformed using a vector as described above.

Yet further the invention provides a method of carrying out gene therapy which comprises forming a vector as described above including a desired gene and incorporating the said vector into the host cells.

By using the chromosome fragments as a basis for DNA vectors, the problems of insertional mutagenesis can be overcome. In particular, they will allow long tracts of DNA, potentially including multiple genes into hosts, suitably mammals or cells in tissue culture. Such vectors, being accurately reproduced during many generations, are extremely advantageous.

As used herein, the expression capable of 'replication and segregation during cell cycle' means that the DNA is accurately replicated and segregated for at least 100 and suitably at least 500 cell divisions and that preferably the chromosome is lost no more than once every 200 cell divisions suitably no more than once every 1000 cell divisions.

In addition, when a fragment is said to be resolved by gel electrophoresis, it is generally expected that this will have a size of no more than about 10Mb. Suitably fragments of this size or less, preferably 6.0Mb or less and more preferably 3.5Mb or less, are used as the basis of DNA vectors.

The invention will now be particularly described by way

of example with reference to the attached Figures in which

Figure 1 shows the design of initial targeting experiment to dissect the centromere region of the human Y chromosome with telomeres;

Figure 2 illustrates the progress of the experiment of Figure 1;

Figure 3 shows the results of an analysis of two cell lines obtained in the experiment by fluorescent *in situ* hybridisation;

Figures 4a-4c show the results of molecular studies carried out on the cell lines of Figure 3;

Figures 5a-5b show the mapping of the EcoRI and Xbal sites flanking the constructs in the cell lines;

Figures 6a-6c show the results of fluorescent *in situ* hybridisation experiments looking at the mitotic behaviour of the Y chromosome and the derivatives;

Figure 7 demonstrates the stable inheritance of one of the cell line and of a deleted derivative of the second;

Figure 8 illustrates the kinetics of variant accumulation in a culture of dividing cells;

Figure 9a and 9b illustrate the design of an experiment to further dissect one of the derivatives and the structure of the plasmid used in this experiment respectively;

Figure 10 illustrates the progress of the second step of the experiment of Figure 9a;

Figure 11 shows a map summary of the desired Δ1 fragment;

Figure 12 illustrates the stability of the final fragment obtained.

Figure 13 shows fluorescent in situ hybridization of chromosomes isolated from the cell line containing Δ1; and

Figure 14 shows pulsed field gel analysis of DNA extracted from the cell line containing Δ1 and from Schizosaccharomyces pombe.

As mentioned above, the applicants have been using cloned telomeric DNA (Barnett et al. 1993) to dissect the human Y chromosome of defined structure (Brown et al. submitted for publication). The design of this work is illustrated in Figure 1a. The 853 mono-chromosome somatic cell hybrid line (Burk et al. 1985) was used as the source of human Y chromosome. This cell line is *hprt*- and will not grow in medium containing hypoxantine, aminopterin and thymidine (HAT). In experiments described in detail below, a plasmid containing a 5.7kb fragment of Y chromosome derived alphoid DNA (Tyler-Smith and Brown, 1987), a gene encoding resistance to the antibiotic G418, a 1.2kb stretch of cloned telomeric DNA (Barnett et al., 1993), and a cassette to allow recovery of flanking sequences in *Saccharomyces cerevisiae* was introduced into 853 cells. Cells which had stably integrated the plasmid were selected and one clone isolated which contained a Y chromosome derivative, called ΔYq74, which had lost the long arm of the human Y chromosome, retained the short arm and which had been truncated by telomere directed breakage within the centromeric array of alphoid DNA. Extensive analysis established that ΔYq74 segregates accurately at mitosis and is retained by cells proliferating for many months in the absence of

selection.

Cells containing ΔYq74 were taken and ΔYq74 broken in a second round of telomere directed breakage. A novel fragment or mini-chromosome designated Δ1 was produced. This was found to be a linear mini-chromosome of about 8Mb in size which could be separated using gel electrophoresis. It is maintained stably by cells proliferating in culture and contains selectable marker genes close to each of the two telomeres.

The following Examples illustrate an embodiment of the invention. In these examples, the following cell culture techniques were employed throughout.

853 cells (Burk et al., 1985), HT1080 cells (Rasheed et al, 1974) and cells containing the truncated chromosomes were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% Foetal Calf Serum as described in the art (Barnett et al., 1993). Electroporation was also as described (Barnett et al., 1993). The protocol for colony hybridization was that of Avraham et al. (1989) except that Hybond nylon membranes were used. Clones to be identified by colony hybridization were originally plated out at about 200 colonies per 9 cm. plate and were subsequently purified by three further rounds of colony hybridization and cloning.

In order to confirm the stability of the clones, the microcell fusion protocol of Fournier (1981) was used except that microcells were separated from donor cells by centrifugation of colcemid treated donor cells for 70 minutes at 35,000g in a 1:1 mixture of Percoll (Pharmacia) and DMEM with 10% Foetal Calf Serum which contained Cytochalsin B at 10μg./ml. Two microcell containing bands formed and were aspirated from the

gradient and the microcells, were harvested by centrifugation at 1,000g for 15 minutes and then filtered successively through 8μm and 5μm polycarbonate filters. Recipient cells were washed twice in phosphate buffered saline, incubated with the microcells for 15 minutes at 37°C in PBS containing Phytohaemagglutinin at 10μg/ml. The supernatant was then gently removed and the cells, coated with the microcells, were incubated in 50% polyethylene glycol 1500 in serum free DMEM for two minutes at 37°C. The cells were washed twice with serum free medium and then incubated for 10-15 minutes at 37°C in serum free medium. The cells were then incubated overnight in DMEM containing FCS, split and selection was applied.

Typically 200μg/ml. G418 was used to select for either HT1080 or 853 derivative cells carrying the G418 resistance gene.

Example 1

Targeting and Cloning of Chromosome Fragments

In an initial targeting experiment (Fig 1) the 853 monochromosomal somatic cell hybrid line (Burk et al., 1985) was used as the source of the human Y chromosome. These cells proliferate quickly in culture and the absence of other human chromosomes in the hybrid make it easy to detect Y chromosome derivatives in a clone of cells. There is a single array of alphoid DNA at the centromere of the human Y chromosome (Tyler-Smith and Brown, 1987). The orientation of the array is known (Cooper et al, 1993) and is indicated by the positions of the EcoRI and XbaI sites which occur once and twice respectively in each 5.7kb unit. Two targeting plasmids (A and B in Fig. 1) were constructed each containing a

5.7kb EcoRI fragment of Y chromosome DNA cloned from the alphoid array (Tyler-Smith and Brown (1987), a gene encoding resistance to the antibiotic G418, a 1.2kb stretch of cloned telomeric DNA (Barnett et al., 1993) and a cassette to allow recovery of flanking sequences in S. cerevisiae.

The G418 resistance gene used in the constructs was 2.2kb Accl-BamHl fragment of pSV_2neo (Mulligan and Berg, 1980). The fragment used in the targeting constructs to allow shuttling into S. cerevisiae was constructed from the URA3 containing Sal-1Xhol fragment of pYAC4 (Burke et al., 1987) the CEN4 containing HpalClal fragment of pYAC4 and an ARS sequence contained in a 500bp SstII-HindIII fragment of the HO endonuclease gene (Kearsey, 1983). The human telomeric sequences were contained in the 1.2kb Taql fragment described in Barnett et al. (1993).

The alphoid DNA in the appropriately linearized plasmid was intended to target it to the alphoid array and the telomeric DNA acts to cause chromosome breakage. The two plasmids differ in the respective orientations of the 5.7kb alphoid units and thus targeted breakage of the Y chromosome by these plasmids was expected to generate derivative chromosomes of reciprocal structures.

The 853 cells were transfected with either of the linearized targeting plasmids A or B and screened in three successive ways for clones which contained Y chromosomes broken within the alphoid DNA. Firstly colony hybridization with Y long arm (DYZ1) and Y short arm (DXYS20) specific probes was used to identify clones which have deleted either all or part of either the long or short arms of the chromosome. Secondly gel electrophoresis and filter hybridization was used to other Y chromosome probes to map the extent of the deletions. Filter hybridization and gel electrophoresis

were carried out as described in Barnett et al., (1993). Finally *in situ* hybridization with a probe specific for the breakage construct was used to determine which clones contained autonomous chromosomes that had been truncated by the targeting constructs. In this way 24 clones were isolated (Fig 2) which were truncated on either the long arm or short arms by telomeric DNA. The sequence content of these clones was determined by gel electrophoresis and filter hybridisation. These are is indicated in Table 1 where 'T' represents truncation of the centromeric alphoid array; 'R' represent a rearrangement characterized by the presence of two cognate large BamHI fragment, '+' indicated that the locus is present and '-' indicates that the locus is absent.

The remaining clones which hybridized differentially to DXYS20 and DYZ1 contained Y chromosome fragments that had either integrated into or translocated onto hamster chromosomes.

Probes used in the analysis of the Y chromosomes were as follows; 29c1 for the telomere of Yp (Cooke et al., 1986), 62RI (Brown, 1988), ZFY (Page et al., 1987), 115i (DXYS8) and p16(DXYS6), (Geldwerth et al., 1986) the 4.5kb HindIII fragment of Y190 (DYZ5; Tyler-Smith et al., 1988) pDP34 (DXYS1) (Page et al., 1984) 50f2, 52d and 64a, (Guellan et al., 1984), 21A1 (DYZ8; Arnemann et al., 1987), GMGY6 and GMGXY3 (Affara et al., 1986) 43/2 and 20/3, oxY1 (Oakey and Tyler-Smith, 1990) and TelBam3.4 for the telomere of Yq (Brown et al., 1990).

In situ hybridisation of biotin labelled probes to metaphase chromosomes was as described by Buckle and Rack (1993).

One member of each of the short arm and of the long arm deletion series had lost either the entire short arm or

long arm and contained a single alphoid DNA array which was smaller than that of the starting chromosome. These chromosomes; ΔYp134 and ΔYq74, were thus candidates for chromosomes which had been targeted and broken in the alphoid array. The results of an analysis of these lines by fluorescent *in situ* hybridization are illustrated in Fig. 3. Probing chromosomes of either the original 853 line, the ΔYq74 clone or the ΔYp134 clone with human genomic DNA detected a single autonomous chromosome. A Y chromosome alphoid DNA probe identified a single locus of hybridization on a chromosome identical in size and appearance to the one identified by the human genomic DNA probe. The alphoid signal was at primary constriction of the Y chromosome and at or close to the end of ΔYq74 or ΔYp134. A probe for the breakage construct, lacking alphoid and telomeric components, specifically stained both ΔYq74 or ΔYp134.

Molecular studies illustrated in Fig 4 confirmed and extended the results of the *in situ* hybridization.

The work giving rise to Fig 4 was carried out as follows:

Fig 4a. The alphoid DNA was truncated and on the same restriction fragment as the targeting construct in ΔYq74 and ΔYp134.

DNA extracted from either 853 cells or from cells containing ΔYq74 or ΔYp134 was restricted with the indicated enzyme, the digests were analysed by pulsed field gel electrophoresis and filter hybridization with a neo probe specific for the targeting construct and then with a Y alphoid probe. Markers are oligomers of λcI857 DNA.

Fig 4b. The targeting constructs have seeded new telomeres in ΔYq74 and ΔYp134. DNA extracted from either

ΔYq74 or ΔYp134 was digested in solution with Bal31 for progressively increasing lengths of time and then with HindIII. Digests were analysed by gel electrophoresis and filter hybridization with a probe specific for the S. cerevisiae CEN4 sequence in each construct.

Fig 4c. The long range organization of the alphoid DNA in ΔYq74 and ΔYp134. The restriction site mapping illustrated in A was extended and sites for the indicated enzymes placed in the DNA flanking the alphoid DNA in the original Y, ΔYq74 and ΔYp134 to produce the map.

DNA from the 853 and ΔYq74 and ΔYp134 cell lines was restricted with six enzymes that do not cut within the alphoid DNA array and analyzed by filter hybridization after gel electrophoresis (Fig 4a) as described by Barnett et al., (1993). Each chromosome contained a single alphoid DNA fragment which was smaller in ΔYq74 and ΔYp134 than in the original cell line. For example the BgIII fragment derived from the 853 line was 550kb long and the fragments derived from the ΔYq74 and ΔYp134 lines were 145kb and 480kb long respectively. The alphoid fragments in the derived lines hybridized to a probe specific for the breakage construct demonstrating that the breakage constructs had integrated into the alphoid DNA (Fig 4a). The telomeric location of the targeting constructs in ΔYq74 and ΔYp134 was confirmed by demonstrating their sensitivity to Bal31 exonuclease digestion (Fig 4b). The restriction site mapping experiments initiated in Fig 4a were extended and positioned the sites for other enzymes in the DNA flanking the alphoid arrays. Restriction site mapping confirmed that the constructs were intact. These maps were combined to produce a map (Fig. 4c) of the DNA flanking array in the original chromosome. The map was identical to the one generated in previous work (Tyler-Smith, 1987) using double digests and thus these results

combine with the deletion analysis to indicate that the truncations have not been accompanied by other rearrangements. The restriction site mapping also established that ΔYq74 and ΔYp134 hold 70kb of DNA of their respective alphoid arrays in common.

Each 5.7kb alphoid unit contains a single EcoRI site and a pair of Xbal sites. These positions of these sites are known and they can be used to orient individual units. In order to determine whether the constructs had integrated into the alphoid DNA by homologous recombination and whether the DNA held in common by ΔYq74 and ΔYp134 was likely to be exclusively alphoid, the EcoRI and Xbal sites flanking the constructs in ΔYq74 and ΔYp134 (Fig. 5) were mapped.

DNA from either ΔYq74 or ΔYp134 was digested to completion with Bglll, partially restricted with either EcoRI or Xbal and analysed by gel electrophoresis and filter hybridization with a neo probe, illustrated in Figure 5a.

Restriction site maps deduced from Fig. 5a are shown in Fig. 5b. There are a number of unresolved bands at about 23kb in the second track of ΔYq74 in A. The number of Xbal sites represented in the unresolved cluster is either three or four and has been arbritarily indicated as four. The alphoid arrays in the top and bottom panels are not aligned.

Analysis of the ΔYq74 DNA established that the construct had not integrated into the genomic DNA by a simple homologous recombination event as it appeared that three 5.7kb alphoid units flanking the integrated construct in the ΔYq74 cell line had inverted with respect to their usual orientation. The ΔYp134 mapping results were more straightforward and were consistent with the construct

Examination of Mitotic Behaviour

Fluorescent *in situ* hybridization to whole anaphase cells was used to examine the mitotic behaviour of the Y chromosome and the truncated derivatives. The techniques used was modified from the protocol developed by Funabiki et al. (1993). The cells were grown for 36-48 hours after plating onto 22mm square coverslips, rinsed in PBS and then simultaneously fixed and permeabilised by a 30min. incubation in 20mM potassium phosphate, 130mM NaCl, 20mM KCI, 10mM EDTA, 2mM $MgCl_2$, 0.1% Triton X-100 and 0.5% glutaraldehyde (EM Grade from Polysciences), pH7.3 at room temperature. The cells on the coverslip over rinsed three times with PBS and incubated twice with 1mg/ml $NaBH_4$ in dH_2O, washed twice in PBS and incubated for 30 minutes in 5% normal goat serum in PBS. The cells were then washed four times in PBS and fixed a second time by incubation in 50mM ethylene glycol bis-(succinimidylsuccinate) in PBS for 30 minutes at 37°C. After washing in PBS the coverslips were incubated in 100µg/ml. RNAase A in PBS at 37°C for one hour and then rinsed again in PBS. The DNA was denatured by a four minute incubation in 0.2M NaOH at room temperature and the coverslips were flooded with PBS. The coverslips were then plunged into ice cold PBS and then hybridized *in situ* with biotinylated probe as described by Buckle and Rack (1993) and finally washed in SSC at 37°C.

Detection was as described by Buckle and Rack (1993) and images were collected on a MRC 600 confocal microscope and transferred to a Titan 3000 workstation for manipulation and display. 3D projections were calculated and displayed using the software described by Highett and colleagues (1993).

Fig 6a shows anaphase segregation of the intact Y chromosome in the 853 cell line, Fig 6b shows segregation of ΔYq74 and Fig 6c shows what we saw most often with ΔYp134. In about 10% of the mitotic cells (see figure legend for numbers) ΔYp134 specifically lagged with respect to the other chromosomes and a particularly extreme example is seen in Fig 6d (a telophase cell). We examined about 1,000 anaphase and early telophase cells of each of the cell lines containing either the original Y, ΔYq74 and ΔYp134 or the chromosome ΔYp6 which is truncated at the tip of the short arm and which serves to control for any effect that the breakage construct may have upon the behaviour of the chromosome (Table 2). No cells were observed in which both Y chromatids segregated to a single pole (2:0 events) and it was thus concluded that each of the four chromosomes was segregating accurately. Although about 10% of the ΔYp134 cells contained a lagging chromatid, all of those seen in the sample 985 cells had separated from their sisters and appeared to be moving to opposite poles. It was estimated that about 10,000 anaphase and telophase ΔYp134 cells were seen in the course of all experiments and in only two of these did it appear possible that the chromatids were not about to end up in different daughter cells. Between 1% and 5% of the cells contain four chromatids segregating in a 1:1 ratio; analysis of metaphase chromosomes established the presence of similar percentages of tetraploid cells with two chromosomes in the lines; suggesting that the cells with four chromatids were tetraploid cells which may have arisen through

occasional endoreduplication events. Between 2% and 5% of the cells did not contain chromosomes detectable by *in situ* hybridization. These cells appear to have arisen from a failure of the probe to enter the cells because the cultures were also analyzed by colony hybridization to total human genomic DNA and found that greater than 99% of the colonies hybridized specifically. Between 0% (ΔYp6) and 1.1% (ΔYp134) of the cells in these lines contained odd numbers of chromatids. These cells may have arisen from either a detection failure or a failure of chromosome replication. Examination of the DAPI stained chromosomes in the same figures revealed no chromatids outside the polar group of hamster suggesting that none of these cells contained inaccurately segregating chromatids.

Given a sample size of 1,000 there is a greater than 95% probability that we would have detected an inaccurate segregation event if such events were occurring at a frequency of greater than 1 in 300 and so we can say at this level of precision that the intact Y chromosome, ΔYp6 and the ΔYq74 and ΔYp134 chromosomes are segregating accurately. These results therefore demonstrate that the ΔYq74 and ΔYp134 have functional centromeres and we shall henceforth refer to them as, respectively, the short arm and long arm acrocentric derivatives of the Y chromosome.

The relative stability of the chromosomes in mitotically proliferating cells was determined by growing cells in the absence of applied selection and measuring the proportion of the cells retaining the chromosomes by colony hybridization. The 853 line was grown for three months and the lines containing the acrocentric chromosomes and the control chromosome (ΔYp6) for three and five months. The lines containing the acrocentric chromosomes and the control chromosomes were also grown in the presence of G418 to confirm the accuracy of the

assays. The Y chromosome and its derivatives were detected by colony lifts and filter hybridisation with either the DYZ1 probe and then total ΔYq74 DNA (853, ΔYp6, ΔYp134) or with the DXYS20 probe and total hybrid DNA. The figures (Table 3) indicate the proportion of colonies which hybridized to the hybrid cell DNA probe but not to the chromosome specific probe. The number of colonies in the two categories are indicated in brackets.

Table 3 indicates that greater than 95% of the cells of each of the lines had retained the chromosomes after three months growth in the absence of applied selection. After five months growth in the absence of the antibiotic 99% of cells retained the ΔYp6 chromosome, 95% of cells retained the long arm acrocentric chromosome and 86% of cells retained the short arm acrocentric chromosome. Fluorescent *in situ* hybridization to chromosomes prepared from these lines at five months confirmed these figures and demonstrated that greater than 95% of the chromosomes isolated from the lines grown in the absence of selection were structurally intact (not shown). If we assume that cells containing the chromosome grow as fast as those without the chromosome then the loss of the short arm acrocentric after five months growth in the absence of selection is consistent with this chromosome being lost once every 1,000 cell division cycles.

The simple model used in the interpretation of the cell culture data is illustrated in Figure 8. In this figure, the filled circles represent type A cells which contain a chromosome and the empty circles correspond to variant B cells which lack a chromosome. It was considered that parental line, termed A, divides with a first order rate k_1, it generates variants called B with a rate constant k_2 which in turn divide with a first order rate given by k_3. For small values of k_2 the following equation should apply:

$dA/dt=k_1A$;

$dB/dt=k_2A+k_3B$

If A_0 is defined as the number of cells at t=0 then this pair of equations has three solutions according to the relative values of k_1 and k_3. The simplest assumption is that the A and B cells divide at the same rate; ie $k_1 = k_3$ and this leads to the solution for small values of B that:

$A=A_0e^{k1t}$ and $B/A=k_2t$.

These expressions were used in extrapolating the rate of chromosome loss per cell division from the results of the experiments where the cells were grown in the absence of selection and then measured the amount of chromosome loss. The rate at which both the 853 cells and the HT1080 cells double is once every day and so the value of $k_1$is taken as $1d^{-1}$.

The assumption of an absence of differences in growth rate is not necessarily valid given the interclonal variations in a typical cell culture and the rate of loss of the short arm acrocentric may thus be less than one loss event per thousand cell divisions.

The acrocentric chromosomes were transferred to the human cell line HT1080 using microcells in order to exclude the possibility that the retention of the chromosomes in the hybrid cells reflected complementation of a recessive defect in the Chinese hamster chromosomal background. HT1080 is of male origin and includes a Y chromosome. The short arm acrocentric chromosome was transferred intact but the long arm acrocentric chromosome deleted sequences during this procedure. Gel electrophoresis, filter hybridization and phosphorimaging was used to

assess chromosome retention after extended proliferation in the presence and absence of G418. Figure 7 shows filter hybridization of BamHI digests with a neo probe. Size markers were multimers of λcl 857 DNA. The alphoid DNA was intact and the chromosomes were maintained after at least four months (HT-ΔYq74) or four and a half months (HT-ΔYp134) growth in the absence of G418. In order to measure the rate of loss of the experimental chromosomes relative to the rate of loss of the Y chromosome in the parental line we reprobed the filter illustrated in Figure 7 with an alphoid probe and determined the relative intensities of the signals due to the host and donor Y chromosomes; comparison of these values before and after growth indicated that the short arm acrocentric ΔYq74 was retained but that about a third of the deleted derivative of the long arm acrocentric had been lost, suggesting a rate of loss of this chromosome of about once every 500 cycles of cell division.

Example 3

Further Telomere-directed Truncation of ΔYq74

The observation that ΔYq74 segregates accurately and is retained by cells for many months in the absence of selection led to further investigation of this line. The overall experiment with respect to ΔYq74 is illustrated in Figure 9a. The terms DXYS20, DYZ1 and alphoid refer to sequences present in the chromosomes which have been indicated diagrammatically. The terms ADE2, svgpt and svneo refer to selectable marker genes present in the breakage constructs. The letters ACU refers to the ARS, CEN4 and URA3 sequences present in the first round breakage construct.

The structure of the plasmid used in the second round of

telomere directed breakage is illustrated in Fig 9b; this plasmid pBSADE2svgpt TEL contains an *Escherischia coli* xanthine-guanine phosphoribosyl transferase gene driven by an SV40 early region promoter (Mulligan and Berg, 1981), a *Saccharomyces cerevisiae* ADE2 gene (Stotz and Linder, 1990) and a 1.2kb stretch of human telomeric DNA which as previously shown (Barnett et al., 1993) efficiently directs chromosome breakage and the formation of a new telomere in mammalian cells.

The progress of the experiment is summarized in Fig. 10. "In situ hybridization" refers to analysis by fluorescent in situ hybridization to metaphase chromosomes as described by Buckle and Rack (1993).

Cells containing ΔYq74 were transfected with the plasmid pBSADE2svgpt TEL which had been linearized so that the telomeric DNA was located at one end of the construct, selected for stably transfected clones using HAT and screened for clones which contained mini-chromosomes generated by a second round of telomere directed chromosome breakage broken in four successive ways. Firstly colony hybridization with the Y short arm probe DXYS20 was used to identify clones which had deleted the centromere distal end of the ΔYq74 chromosome. Secondly fluorescence *in situ* hybridization with a Y chromosome alphoid DNA probe was used to identify clones containing small autonomous chromosome derivatives of the original ΔYq74 chromosome; this led to the identification of 16 chromosomes. Thirdly gel electrophoresis and filter hybridization to other Y chromosome short arm probes was used to map the extents of the deletions and to detect any rearrangements in this set of 16 clones (Table 4). Finally fluorescent *in situ* hybridization to metaphase chromosomes with a probe for the breakage construct was used to determine which clones contained autonomous chromosomes that had been truncated by the second round

breakage construct. In this way we isolated one clone which contained a chromosome Δ1, which was truncated by the integration of the second round breakage construct pBSADE2svgpt TEL into ΔYq74. The remaining clones contained Y chromosome fragments rearranged by other undefined mechanisms. A final FISH analysis of the clone containing the Δ1 chromosome was carried out (see Fig 13). Metaphase chromosomes of the hybrid somatic cell line containing Δ1 were hybridized in situ with a probe recognising the svgpt gene, the SVneo gene and the S. cerevisiae components in the breakage constructs used to generate Δ1. The probe was not specific for the second round breakage construct, pBSADE2svgpt TEL, but also recognized sequences used in the construction of ΔYq74. As shown in Fig 13, three sites of hybridization were seen, two are on the mini-chromosome, Δ1 and are represented by three dots, indicated with an arrow head, and one is at the end of a large hamster chromosome. The starting chromosome ΔYq74 contains a single site of hybridization with this probe. It was concluded that the line containing Δ1 contains two sites where the pBSADE2svgpt TEL has integrated; one is on the Y derived mini-chromosome Δ1 and the other is on a hamster chromosome. Analysis of DNA extracted from the line containing Δ1 by restriction enzyme digestion and filter hybridization confirmed that there are two sites of integration of pBSADE2svgpt TEL and that they are both telomeric.

Example 4

Size fractionation and stability of Δ1 by pulsed field gel electrophoresis.

Pulsed field gel electrophoresis was used to measure the size of Δ1. We size fractionated DNA extracted from cells containing Δ1 together with DNA extracted from the yeast *Schizosaccharomyces pombe* and analysed the size

fractionated DNA either by ethidium bromide staining and photography (Figure 14A) or by hybridization with a Y chromosome alphoid probe after filter transfer (Figure 14B). DNA embedded in agarose plugs was extracted from either Schizosaccaromyces pombe cells or from cells containing Δ1 and electrophoresed at 1 V cm^{-1} in 0.6% agarose in a pulsed field gel in half strength TAE buffer at a pulse time of 90 minutes for 10 days at 4 degrees celsius. The gel was then stained in a solution of ethidium bromide at 10 μg/mL for 1h, rinsed in water and photographed under ultra violet light, transferred to a nylon filter and then the filter was probed with a Y chromosome alphoid DNA probe. An autoradiograph of the hybridized filter transfer was made.

The relative mobilities of the hybridizing fragment in Figure 14B and of the *S. pombe* chromosomes which are of known size (Fan et al, 1988) indicated that the Δ1 chromosome is approximately 8Mb in size.

The results are summarized in the map of Δ1 illustrated in Fig 11. As noted earlier the human Y chromosome has been extensively mapped (Foote et al., 1992) and it is possible to estimate the approximate positions on Δ1 of some of the sequences used in the analysis of Δ1 indicated in Table 1.

Pulsed field gel electrophoresis was also used to examine the structural stability of Δ1 upon prolonged proliferation. Cells were cultured for three months and DNA was extracted at intervals of approximately 14 days as illustrated in Fig 12. The DNA was size fractionated by gel electrophoresis, filter transferred to nylon and analysed by hybridization with a probe specific for the alphoid DNA; Δ1 is maintained with a constant structure.

ΔYq74 and Δ1 were analysed for sequence tagged sites

(STS) as set out in Vollrath et al., Science (1992), 258. 52-59 which is incorporated herein by reference, using the polymerase chain reaction (PCR). It was found that ΔYq74 contains sY 3, 5, 6, 7,10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 76 and 77. Δ1 retains all these sitses with the exception of the more distal sequences, specifically sY 3, 5, 6, 7, 10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, and 33.

Example 5

Further Telomeric-directed Truncation of ΔYq134

Using methods anologous to those described in Example 3 and Example 4, three stable derivatives of the ΔYq134 clone were obtained and designated Δ7, Δ128 and Δ196. Δ7 and Δ128 are similar but not identical in size and are approximately 6Mb long. Δ196 is approximately 9Mb long.

Analysis of ΔYq134, Δ 7, Δ128 and Δ196 for the long arm sequence tagged sites have found that ΔYq134 contains sY 53, 54, 55, 56, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 90, 92, 93, 94, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 113, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 157, 158, 159, 160 and 161.

Δ 7 was analysed for all these sites but it is deleted for all except sY 54, 55, 56, 88, 90, 92, 97, 98, 99, 106, 109, 110, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 142, 143, 150, 153, 154, 155, 159,

160 and 161 which are retained.

Similarly Δ128 was analysed for all the sequences found on ΔYq134 and it is deleted for all except sY 54, 55, 56, 88, 90, 92, 97, 98, 99, 104, 106, 109, 110, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 142, 143, 150, 153, 154, 155, 159, 160 and 161 which are retained.

Analysis of Δ196 in a similar manner showed that all the STS sequences present on ΔYq134 were deleted with the exception of sY 79, 81, 82, 83, 84, 85, 86, 87, 88, 90, 92, 93, 94, 97, 98, 99, 100, 101, 102, 151 and 161.

All chromosome fragments retain the alphoid DNA and are therefore positive for sY78.

Example 6
Vector production

Δ1 is linear mini-chromosome of approximately 8Mb in size; it is maintained stably by cells proliferating in culture and contains selectable marker genes close to each of the two telomeres. Δ1 can be readily engineered to act as a vector for specific sequences by using homologous recombination to target sequences into the region of Δ1 containing the svgpt gene. Sites (eg the lox P site of phage P1) for the site specific recombinases (eg the cre recombinase of phage P1) could be targeted to Δ1 as described above and used as targets for recombinase catalysed integration of specific sequences of interest such as plasmids or yeast artificial chromosomes. This would be a specific example of a general use of such site specific recombinases proposed by other (Fukushige and Sauer, 1992). The recombinase could also be used to catalyse the translocation of sequences between the proposed

derivative of Δ1 and host cell chromosomes which have also been marked with a site specific recombination recognition sequence. In this way large tracts of DNA could be mobilized onto the derivative of Δ1.

Similar considerations will apply to Δ7, Δ128 and Δ196 in terms of their use as vectors.

Several human genes or genes complexes of clinical importance are located close to telomeres and could be mobilized onto the derivative of Δ1 in this way. These genes include the α globin complex, the factor VIII gene and the immunoglobulin heavy chain complex.

Δ1 and derivatives of the types described above could be used in transgenesis of whole animals by introducing these mini-chromosomes into embryonal stem cells using either fusion or microcells (Fournier, 1981) containing Δ1 or a derivative with embryonal stem cells. Alternatively Δ1 or its derivatives might be isolated from cells arrested at mitosis and introduced into embryonal stem cells by electroporation or lipofection (Felgner et al., 1987). In those organisms where embryonal stem cell lines have not been isolated Δ1 or its derivatives be introduced into fertilized eggs by microinjection.

Similarly Δ1 or its derivatives could be used as vectors for human gene therapy by introducing the mini-chromosomes into human somatic cell types such as, but not exclusively, hepatocytes, keratinocytes, bone-marrow cells or blood cells by microcell fusion, electroporation or lipofection.

Although the techniques have been applied with respect to the human Y chromosome, the skilled person would understand that other chromosomes, in particular plant

chromosomes, could be subjected to similar analysis techniques and form the basis of related vectors. For example, it is known from the study of breaks induced in chromosomes by X-rays that it is possible to subdivide plant centromeres (McClintock, 1932). Subsequent studies have supported this observation (Rhodes, 1938; Darlington, 1940) and led to the idea that functional entities of centromeres were repeated. This means that fragments of the required size and with the required centromeric function could be obtained from plant chromosomes using similar reaction strategies to those described above.

References

Affara et al (1986), "Regional Assignment of Y linked DNA probes by deletion mapping and their homology with X-chromosome and autosomal sequences", Nucl. Acid Res., 14, 5353-5373.

Arnemann et al (1985), "Human Y chromosome derived cloned DNA sequences", Cytogenet. Cell Genet., 40, 571.

Avraham et al, (1989), "A Rapid Procedure for Cell Colony Hybridization using DNA probes", Anal. Biochem., 179, 217-221.

Barnett et al, (1993), "Telomere directed fragmentation of Mammalian Chromosomes", Nucl. Acid Res. 21, 27-36.

Brown, W.R.A., (1988), "A physical map of the human pseudoautosomal region", EMBO J., 7, 2377-2385.

Brown, W.R.A., (1992), "Mammalian artificial chromosomes", Current Opinion in Genetics and Development, 2, 479-486.

Brown et al, (1990), "Structure and polymorphism of Human telomere-associated DNA", Cell 63, 119-132.

Buckle et al, (1993), "Fluorescent in situ hybridization in Human Genetic Disease Analysis", (2nd Edition); ed K.E. Davies, IRL Press.

Burk et al, (1985), "Characterization of a single copy sequence from the human Y chromosome", Mol. Cell. Biol., 5, 576-581.

Burke et al, (1987), "Cloning of large segments of exogenous DNA by means of artificial chromosome vectors",

Science 236, 806-812.

Cooke, H.J., (1976), "A repeated sequence specific for human males", Nature, 262, 182-186.

Cooke et al, (1985), "Hypervariable telomeric sequences from the human sex chromosomes are pseudoautosomal", Nature, 317, 687-692.

Cooper et al, (1992), "The putative centromere forming sequence CM8 is a single copy sequence and not a component of most centromeres", Hum. Molec. Gen., 1, 753-754.

Cooper et al (1993a), "Structure of the sequences adjacent to the centromeric alphoid satellite DNA on the human Y chromosome", J. Mol. Biol., 230, 787-799.

Cooper et al, (1993b), "The major centromeric array of alphoid satellite DNA on the human Y chromosome is non-palindromic", Human. Molec. Gen., 2, 1267-1270.

Darlington, C.D. (1940), "The origin of Iso-Chromosomes", J. Genet., 39, 351-361.

Fan et al, (1988), Nucleic Acids Res., 17, 2801-2818.

Felgner et al, (1987), "Lipofection: A highly efficient, lipid mediated DNA-transfection procedure", Proc. Natl. Acad. Sci. (USA), 84, 7413-7417.

Foote et al, (1992), "The human Y chromosome: overlapping clones spanning the euchromatic region", Science, 258, 60-66.

Fournier, R.E.K. (1981), "A general high-efficiency procedure for production of microcell hybrids", Proc.

Natl. Acad. Sci. (USA), 78, 6349-6353.

Funabiki et al, (1993), "Cell cycle dependent specific positioning and clustering of centromeres and telomeres in Fission Yeast", J. Cell. Biol., 121, 961-976.

Fukushige et al, (1992), "Genomic targeting with a positive-selection/ox integration vector allows highly reproducible gene expression in mammalian cells", Proc. Natl. Acad. Sci. (USA), 89, 7905-7909.

Geldwerth et al, (1985), "Extensive DNA sequence homologies between the human Y chromosome and the long arm of the X chromosome", EMBO J., 4, 1739-1743.

Guellan et al, (1984), "Human XX males with Y single copy DNA fragments", Nature, 307, 172-173.

Highett et al, (1993), "Different Patterns of rDNA distribution in Pisum sativum correlate with different levels of nuclear activity", J. Cell Sci., 104, 843-852.

Guellan et al, (1984), "Human XX males with Y single copy DNA fragments", Nature, 307, 172-173.

Haaf et al, (1992), "Integration of human α satellite DNA into simian chromosomes; centromere protein binding and disruption of normal chromosome segregation", Cell, 70, 681-696.

Hadlaczky et al, (1991), "Centromere formation in mouse cells cotransformed with human DNA and a dominant selectable marker", Proc. Natl. Acad. Sci., (USA), 88, 8106-8110.

Heartlein et al, (1988), "Chromosome instability associated with human alphoid DNA transfected into the

Chinese hamster genome", Mol. Cell. Biol. 8, 3611-3618.

Kearsey, S., (1983), "Analysis of sequences conferring autonomous replication in Baker's yeast", EMBO J., 2, 1571-1575.

Manuelidis, L., (1978), "Chromosomal localization of complex and repeated human DNA", Chromosoma, 66, 23-32.

McClintock, B., (1932), "A correlation of ring-shaped chromosomes with variegation in Zea Mays", Proc. Natl. Acad. Sci. (USA), 18, 677-681.

Mulligan et al, (1980), "Expression of a bacterial gene in mammalian cells", Science, 209, 1422-1428.

Mulligan et al, (1981), "Selection for animal cells that express the *Escherichia coli* gene coding for xanthine-guanine phosphoribosyltransferase", Proc. Natl. Acad. Sci. (USA), 78, 2072-2076.

Murray et al, (1993), "The Cell Cycle"; Oxford Univ. Press.

Nurse, P., (1980), "Universal control mechanism regulating the onset of M phase", Nature, 344, 503-508.

Oakey et al, (1990), "Y chromosome haplotyping suggests that most European and Asian men are descended from one or two males", Genomics, 7, 325-330.

Page et al, (1984), "Occurrence of a transposition from the X-chromosome long arm to the Y chromosome short arm during human evolution", Nature 311, 119-123.

Page et al, "The sex determining region of the human Y chromosome encodes a finger protein", Cell, 51, 1091-

1104.

Rasheed et al, (1974), "Characterization of a newly derived human sarcoma line (HT1080)", Cancer, 33, 1027-1033.

Rhoades, M.M., (1938), "On the origin of secondary trisome through the doubling of a half-chromosome fragment", Genetics, 23, 163-164.

Rouyer et al, (1986), "A gradient of sex linkage in the pseudoautosomal region of the human sex chromosomes", Nature 319, 291-295.

Smith G.P., (1976), "Evolution of repeated DNA sequences by unequal cross-over", Science, 191, 528-535.

Southern, E.M., (1975), "Long range periodicities in mouse satellite DNA", J. Mol. Biol., 94, 51-69.

Stotz et al, (1990), "The ADE2 gene from *Saccharomyces cerevisiae*, sequence and new vectors", Gene, 95, 91-98.

Surosky et al, (1985), "Construction of telocentric chromosomes in Saccharomyces cerevisiae", Proc. Natl. Acad. Sci. (USA)., 82, 2106-2110.

Tyler-Smith, C., (1987), "Structure of repeated sequences in the centromeric region of the human Y chromosome", Differentiation (Suppl.), 101, 93-100.

Tyler-Smith et al, (1987), "Structure of the major block of alphoid satellite DNA on the human Y chromosome", J. Mol. Biol., 195, 457-470.

Tyler-Smith et al, (1988), "Structure of a hypervariable tandemly repeated DNA sequence on the short arm of the

human Y chromosome", J. Mol. Biol., 203, 837-848.

Tyler-Smith et al, (1993), "Localization of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes", Nature Genet., 5, 368-375.

Claims

1. A vector comprising a chromosome fragment which fragment comprises sequences which have centromeric activity, is capable of replication and segregation during cell cycle, and is of such a size that it can be resolved using gel electrophoresis.

2. A vector according to claim 1 which is derived from the human Y chromosome.

3. A vector according to claim 2 comprising a linear derivative of the human Y chromosome which includes at least some alphoid DNA.

4. A vector according to claim 2 or claim 3 which is derived solely from the short arm of the human Y chromosome.

5. A vector according to claims 2 or claim 3 which is derived solely from the long arm of the human Y chromosome.

6. A vector according to any one of the preceding claims which is less than 6Mb in size.

7. A vector according to claim 6 which is less than 3.5 megabase in size.

8. A process for obtaining a vector according to any one of the preceding claims which process comprises dissecting a chromosome in such a manner that a fragment retaining centromere function is retained.

experiments and only those which are mitotically stable to a reasonable extent selected for further truncation.

10. A process according to claim 9 wherein the said mitosis experiments are carried out using *in situ* hybridisation to chromosomes in whole cells.

11. A process for obtaining a vector which process comprises dissecting a chromosome using a pair of vectors each containing telomeric DNA, a marker gene, and chromosomal DNA from the centromeric region in opposite orientations in the two vectors, introducing the vectors into cells containing the chromosome, selecting cells which have stably integrated vector, and isolating a clone containing a chromosome derivative which has lost one arm of the chromosome but retained the other arm truncated by telomere-directed breakage within the centromeric array, such clone segregating accurately at mitosis and being retained by cells proliferating in the absence of selection.

12. A process according to claim 11 wherein the selected clone is subjected to a further round of telomere-directed breakage to obtain a mini-chromosome fragment which is maintained stably by cells proliferating in

culture.

13. A process according to claim 11 or claim 12 which comprises the steps of carrying out a first telomere-directed truncation of the human Y chromosome, isolating stably segregating a fragment which contains only short arm and alphoid DNA, carrying out a second telomere-directed truncation of the isolated fragment, and isolating a shorter stably segregating fragment.

14. A vector according to any one of claims 1- 7 which further comprises a gene encoding a desired polypeptide or protein.

15. A cell transformed using a vector according to claim 14.

16. A vector according to claim 14 for use in gene therapy.

17. A pharmaceutical composition comprising a vector according to claim 16 in combination with a pharmaceutically acceptable carrier or diluent.

18. A vector according to claim 14 for use in animal transgenesis.

19. A method for treating a host by administering a gene which expresses a polypeptide or protein of therapeutic use to said host, which method comprises forming a vector as described above including a desired gene and incorporating the said vector into the host cells.

20. A vector obtainable by a process according to claim 11.

21. A vector comprising Δ1 obtainable from cell line

having the accession number ECACC 95052342.

22. A vector comprising Δ7 obtainable from cell line having the accession number ECACC 95052343.

23. A vector comprising Δ128 obtainable from cell line having the accession number ECACC 95052344.

24. A vector comprising Δ196 obtainable from cell line having the accession number ECACC 95052345

25. A cell line selected from any one of those having the accession numbers ECACC95052342, ECACC 95052343, ECACC 95052344 and ECACC 95052345.

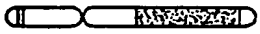
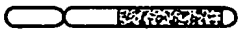
Cell lines		Tel	XYS20	62	ZFY	XYS8	XYS6	YZ5'	XYS1	50f2B	50f2A	YZ5	50f2d	α	64A	Y6	XY3	43/2	52dA	50f2E	YZ8	50f2C	20/3	oxYI	YZI	Tel
ΔYp	134	-	-	-	-	-	-	-	-	-	-	-	-	T	+	+	+	+	+	+	+	+	+	+	+	+
	38	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	23	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	122	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	102	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	29,149	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	6,51,87,120,126	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ΔYq	54,136	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
	98,104	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-
	22,148	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-
	115	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-
	74	+	+	+	+	+	+	+	+	+	+	+	+	T	-	-	-	-	-	-	-	-	-	-	-	-
	112	+	+	+	+	+	+	+	+	+	+	+	+	R	+	+	-	-	-	-	-	-	-	-	-	-
	69	+	+	+	+	+	+	+	+	+	+	+	+	R	+	-	-	-	-	-	-	-	-	-	-	-
	110	+	+	+	+	+	+	+	+	+	-	-	-	R	+	-	-	-	-	-	-	-	-	-	-	-
	130	+	+	+	+	+	+	+	+	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-	-	-

TABLE 1

Chromosome	1:1	2:2	2:0	1:0	0:0	2:1	Other
Y	965(0.96)	16(0.016)	0	0	22(0.022)	1(0.001)	2(0.002)
ΔYp6	950(0.94)	11(0.01)	0	0	52(0.051)	0	1(0.001)
ΔYq74	1194(0.91)	52(0.04)	0	0	51(0.04)	5(0.004)	6(0.005)
ΔYp134	1061(0.88)	100(0.082)	0	4(0.003)	26(0.021)	10(0.008)	11(0.008)

TABLE 2

Three months

Clone		-G418	+G418
	853/4	0.006 (7/1118)	/
	Δ Yp 6	0.01(11/1083)	0.004(5/1312)
	ΔYq 74	0.04 (48/1127)	0 (0/802)
	Δ Yp 134	0.02(21/1039)	0 (0/1296)

Five months

Clone		-G418	+G418
	Δ Yp 6	0.006(16/2563)	0.037(59/1594)
	Δ Yq 74	0.14(136/987)	0.005(10/2029)
	Δ Yp 134	0.05(48/961)	0.002(2/1237)

TABLE 3

Cell lines	Locus												
	XYS20	XYS17	ZFY	XYS8	XYS6	YZ5'	XYS1	50f2B	50f2A	YZ5	50f2d	α
37	-	+	+	+	+	r	+	+	-	r	+	+
3,24a,27,30,31	-	+	+	+	+	r	+	+	+	r	+	r
15a	-	+	+	+	+	+	+	+	+	+	+	r
17b	-	-	+	+	+	r	+	+	+	r	+	r
24b	-	-	-	-	-	-	-	-	r	+	+	r
4	-	-	+	+	+	+	+	+	+	+	+	+
Δ1,39	-	-	-	+	+	+	+	+	+	+	+	+
21b	-	-	-	-	-	+	+	+	+	+	+	+
28	-	-	-	-	-	-	+	+	+	+	+	+
21a	-	-	-	-	-	-	-	-	+	+	+	+
17c	-	-	-	-	-	-	-	-	-	+	+	+

TABLE 4

Fig.1.

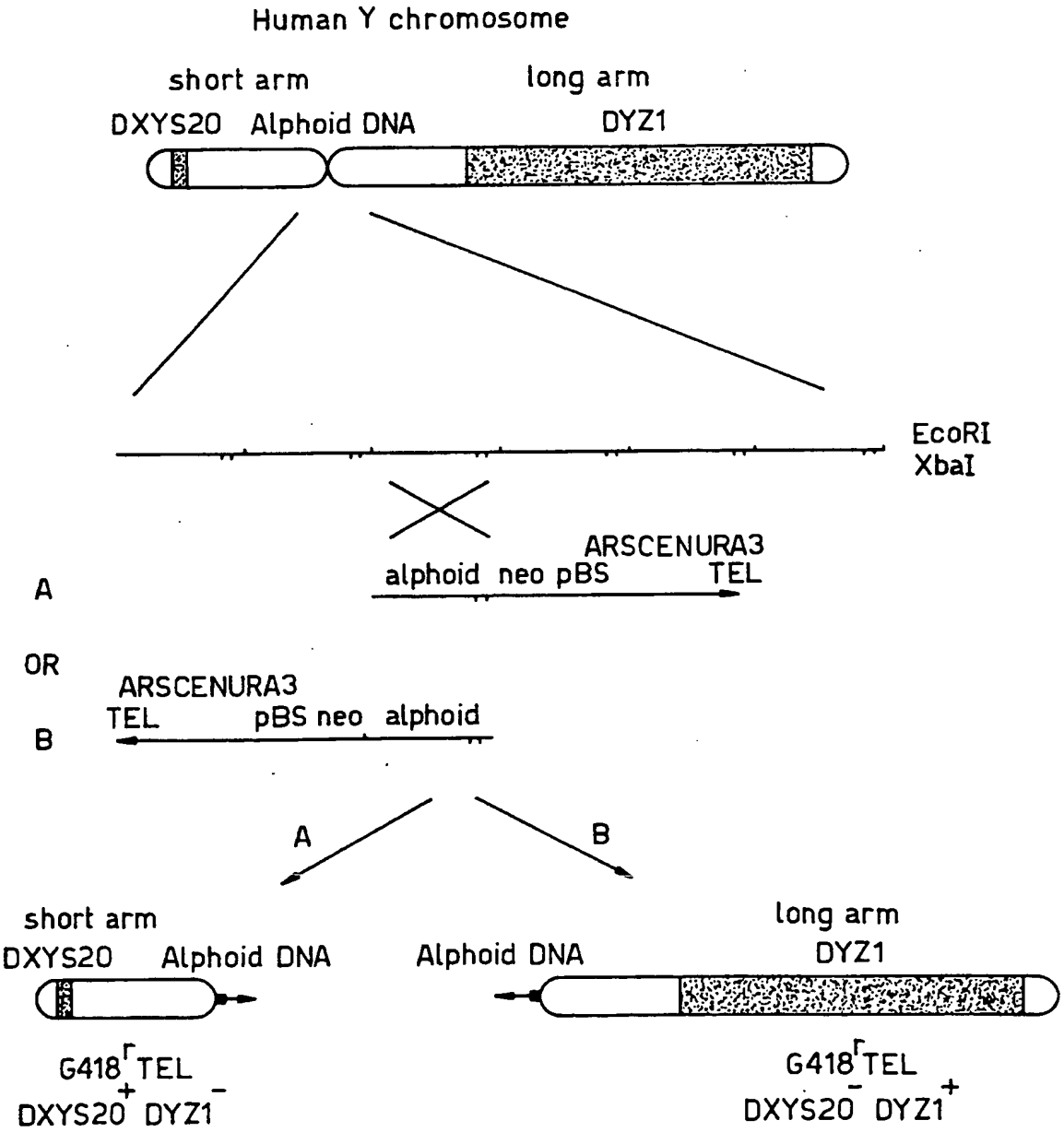

Human Y chromosome
short arm
long arm
DXYS20
Alphoid DNA
DYZ1
EcoRI
XbaI
ARSCENURA3
alphoid neo pBS
TEL
A
OR
B
ARSCENURA3
TEL
pBS neo alphoid
A
B
short arm
DXYS20
Alphoid DNA
long arm
Alphoid DNA
DYZ1
G418r TEL
DXYS20+ DYZ1-
G418r TEL
DXYS20- DYZ1+

Fig.2.

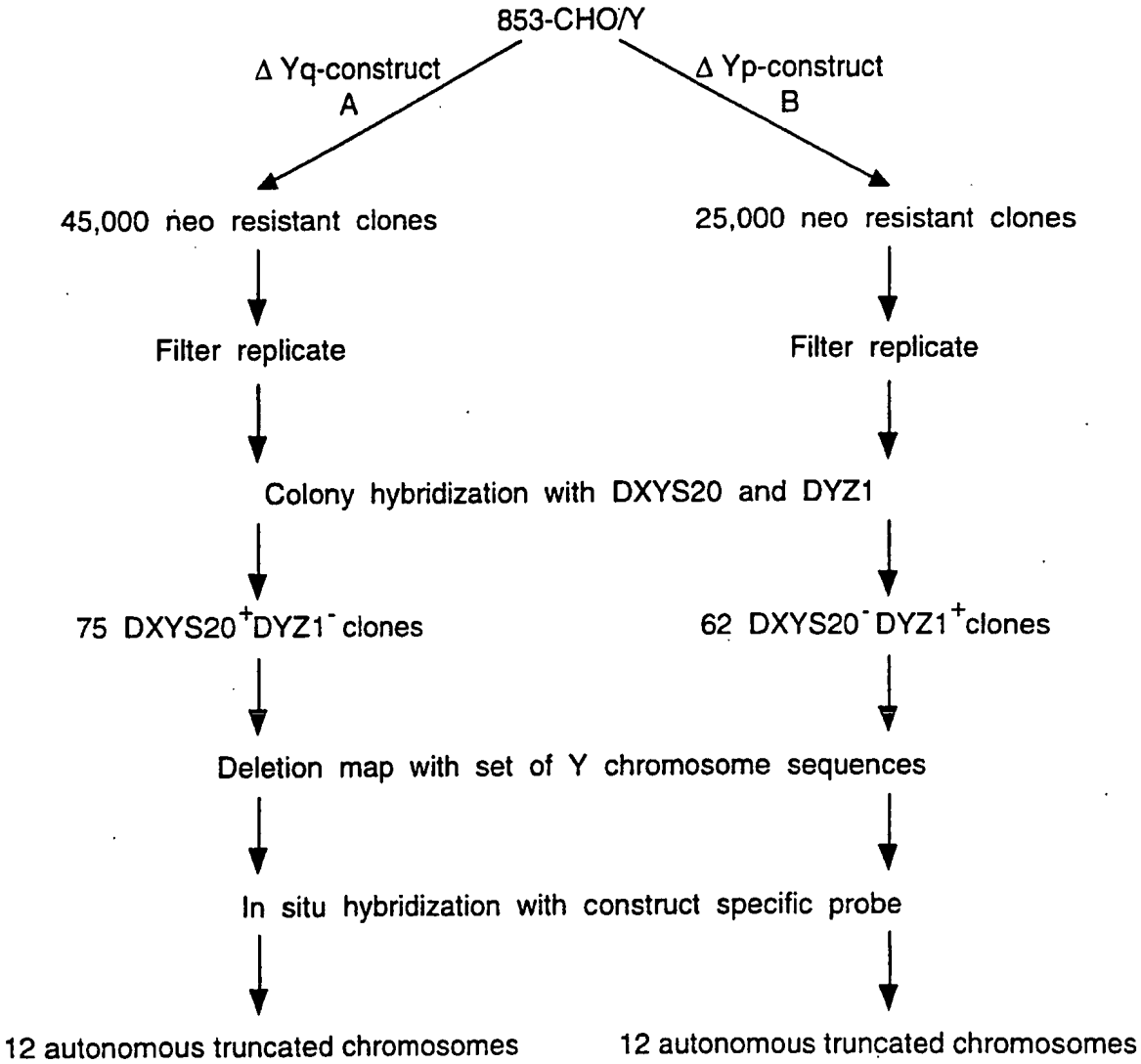

853-CHO/Y
Δ Yq-construct
A
Δ Yp-construct
B
45,000 neo resistant clones
25,000 neo resistant clones
Filter replicate
Filter replicate
Colony hybridization with DXYS20 and DYZ1
75 DXYS20+DYZ1- clones
62 DXYS20-DYZ1+clones
Deletion map with set of Y chromosome sequences
In situ hybridization with construct specific probe
12 autonomous truncated chromosomes
12 autonomous truncated chromosomes

Fig.3.

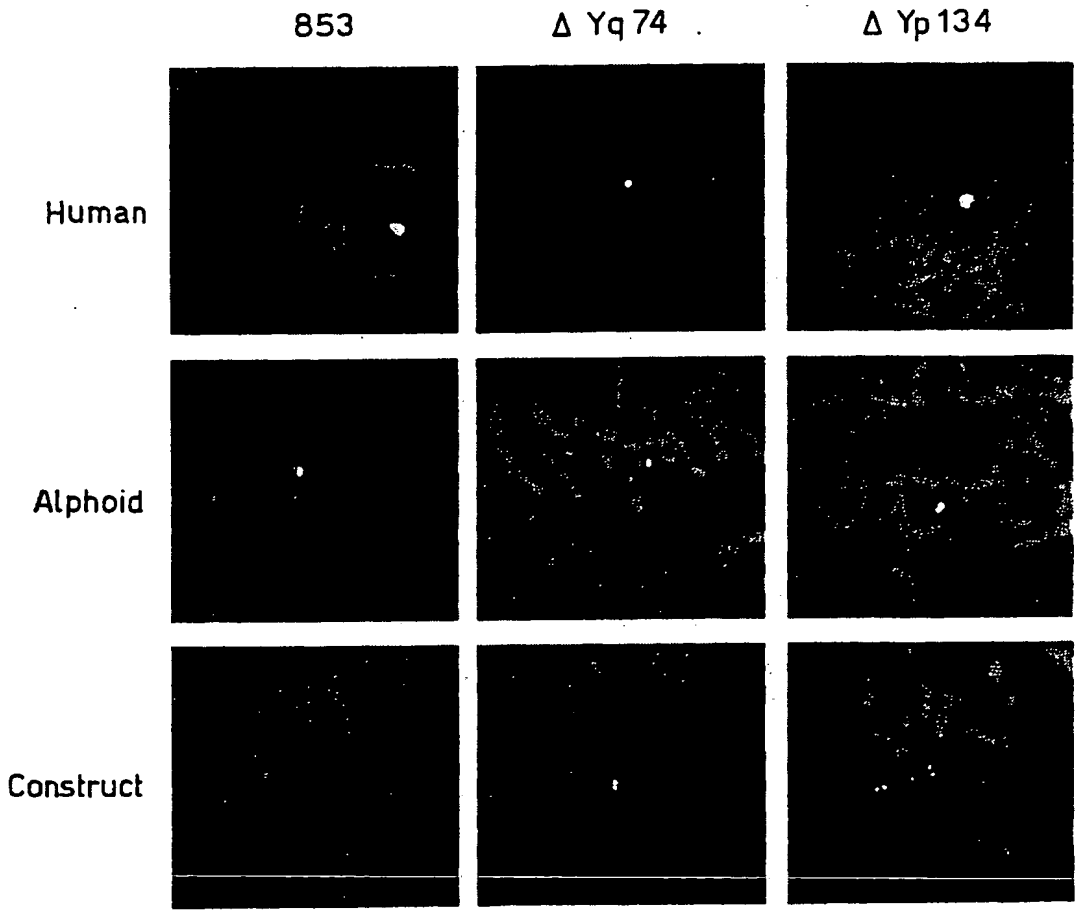


Fig.4a.

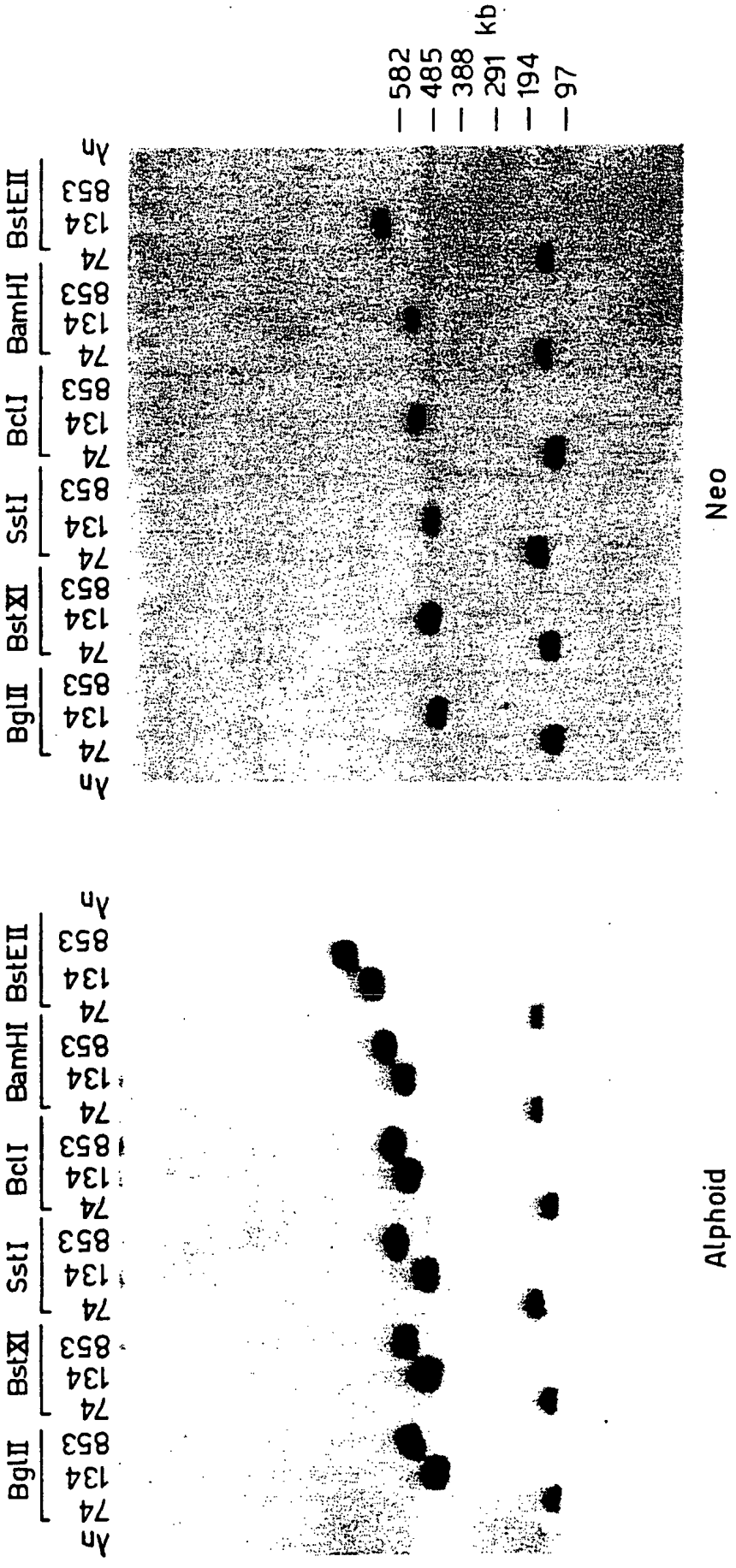

kb
582
485
388
291
194
97
λn
74
134
853
BglII
74
134
853
BstXI
74
134
853
SstI
74
134
853
BclI
74
134
853
BamHI
74
134
853
BstEII
λn
Neo
Alphoid

Fig.4b.

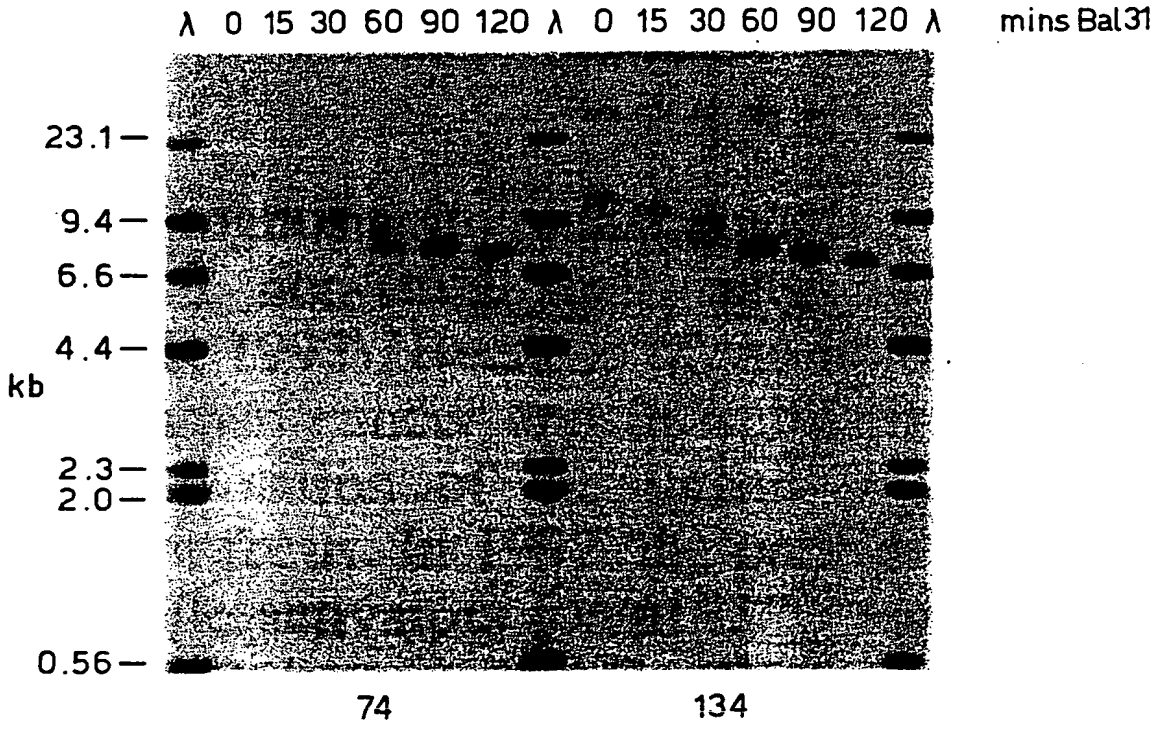

λ 0 15 30 60 90 120 λ 0 15 30 60 90 120 λ
mins Bal31
23.1
9.4
6.6
4.4
kb
2.3
2.0
0.56
74
134

Fig.4c.

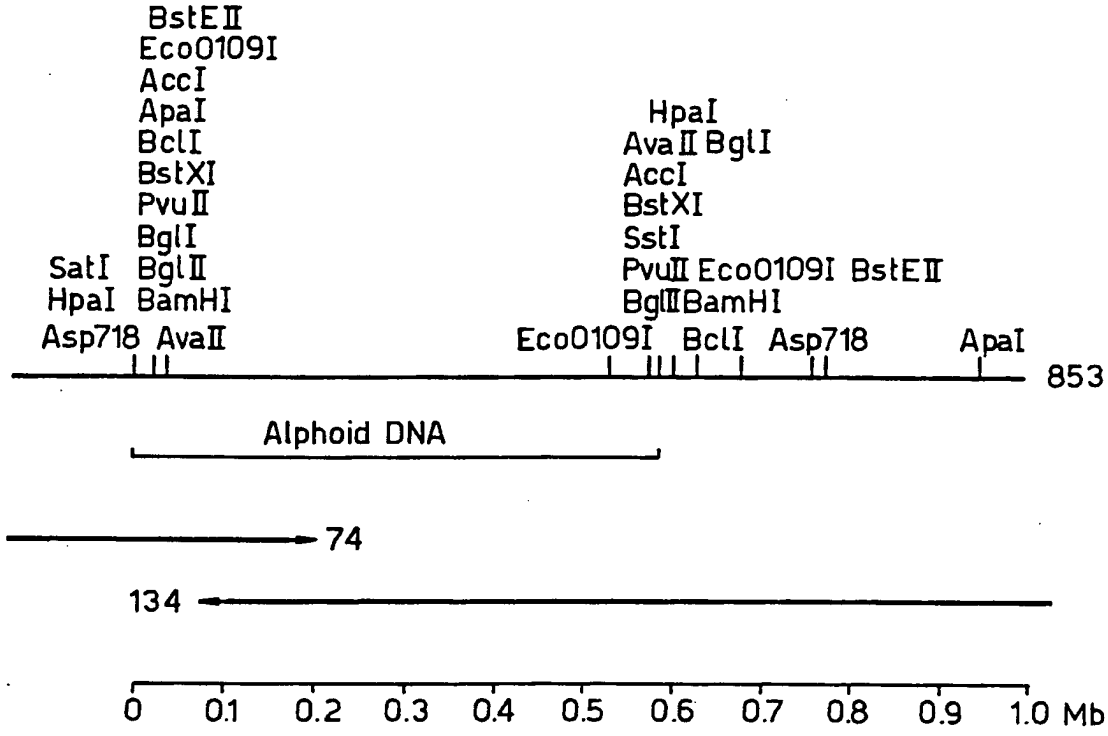

BstEII
Eco0109I
AccI
ApaI
BclI
BstXI
PvuII
BglI
SatI BglII
HpaI BamHI
Asp718 AvaII
HpaI
AvaII BglI
AccI
BstXI
SstI
PvuII Eco0109I BstEII
BglII BamHI
Eco0109I BclI Asp718 ApaI
853
Alphoid DNA
74
134
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 Mb

Fig.11.

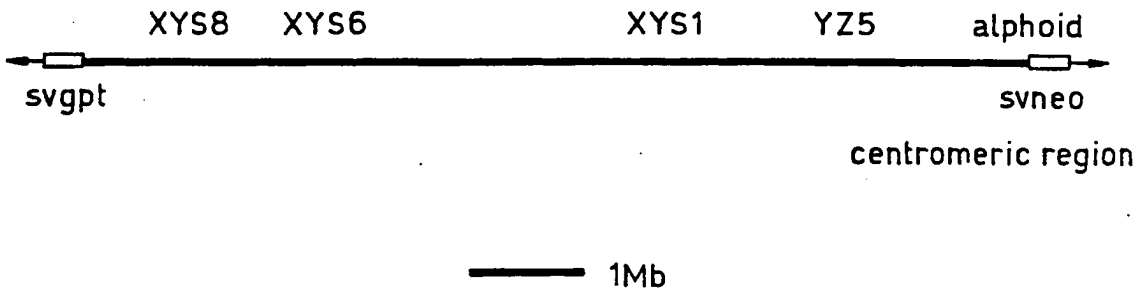

XYS8 XYS6 XYS1 YZ5 alphoid
svgpt
svneo
centromeric region
1Mb

Fig.5a.

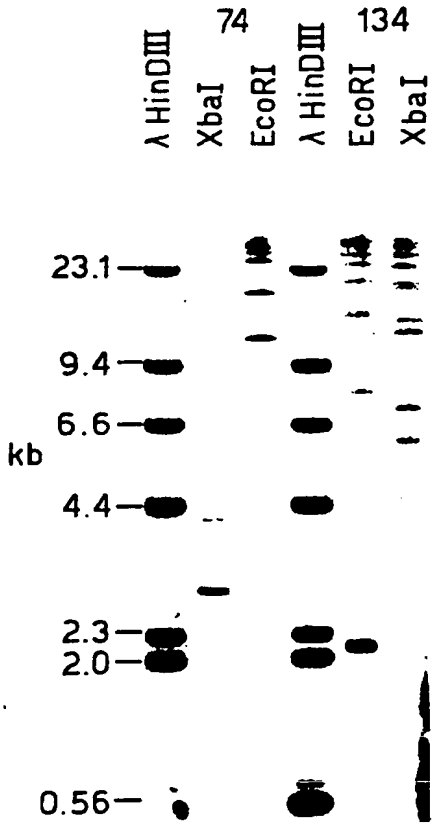

λ HinDIII
XbaI
74
EcoRI
λ HinDIII
134
EcoRI
XbaI
23.1
9.4
6.6
kb
4.4
2.3
2.0
0.56

Fig.5b.

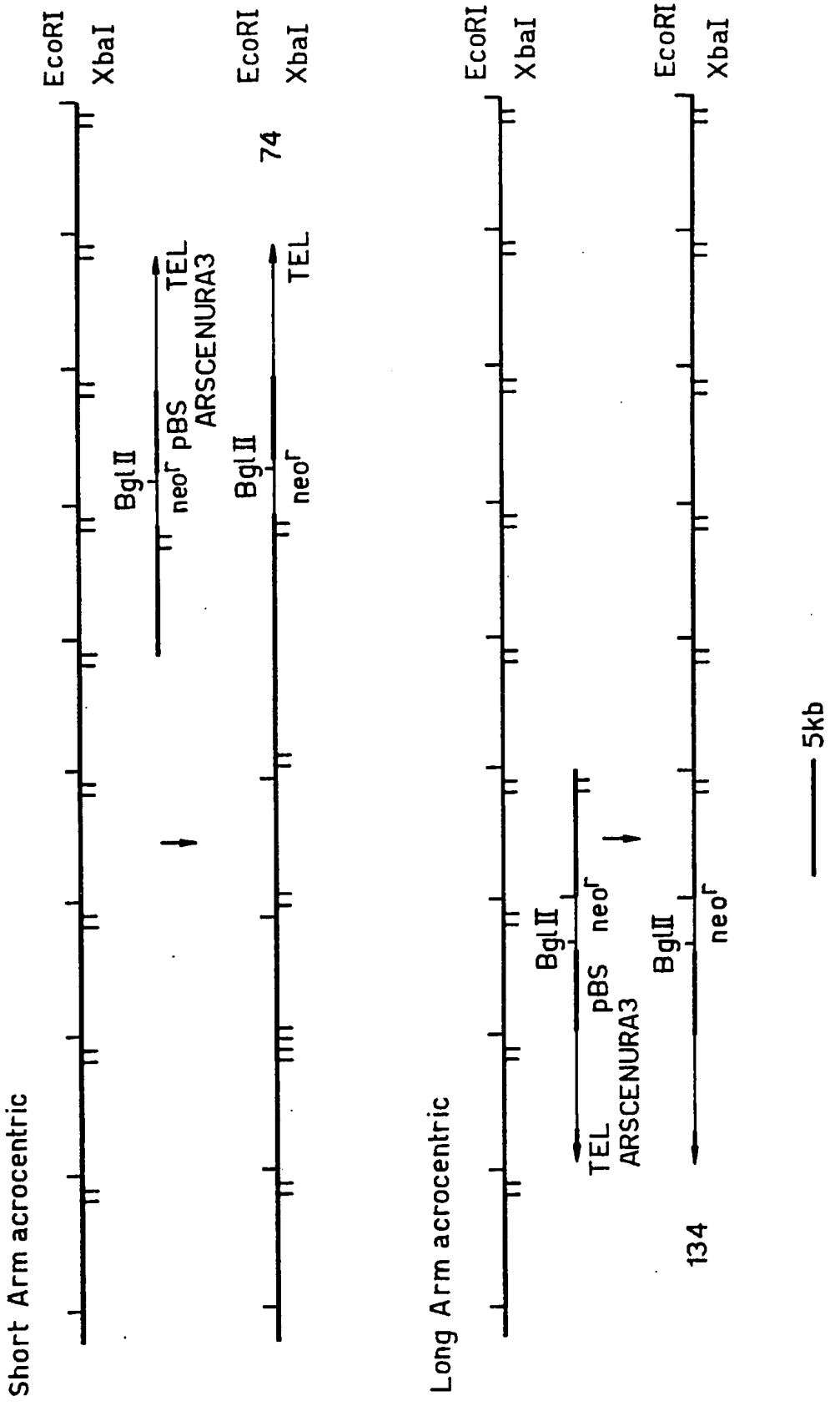
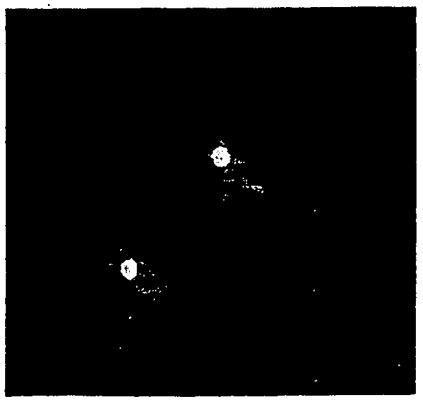

Short Arm acrocentric
EcoRI
XbaI
BglII
neo^r
pBS
ARSCENURA3
TEL
74
Long Arm acrocentric
134
5kb

Fig.6.



A: Intact Y chromosome

B: ΔYq 74

C: ΔYp134

D: ΔYp134-lagging

Fig.7.

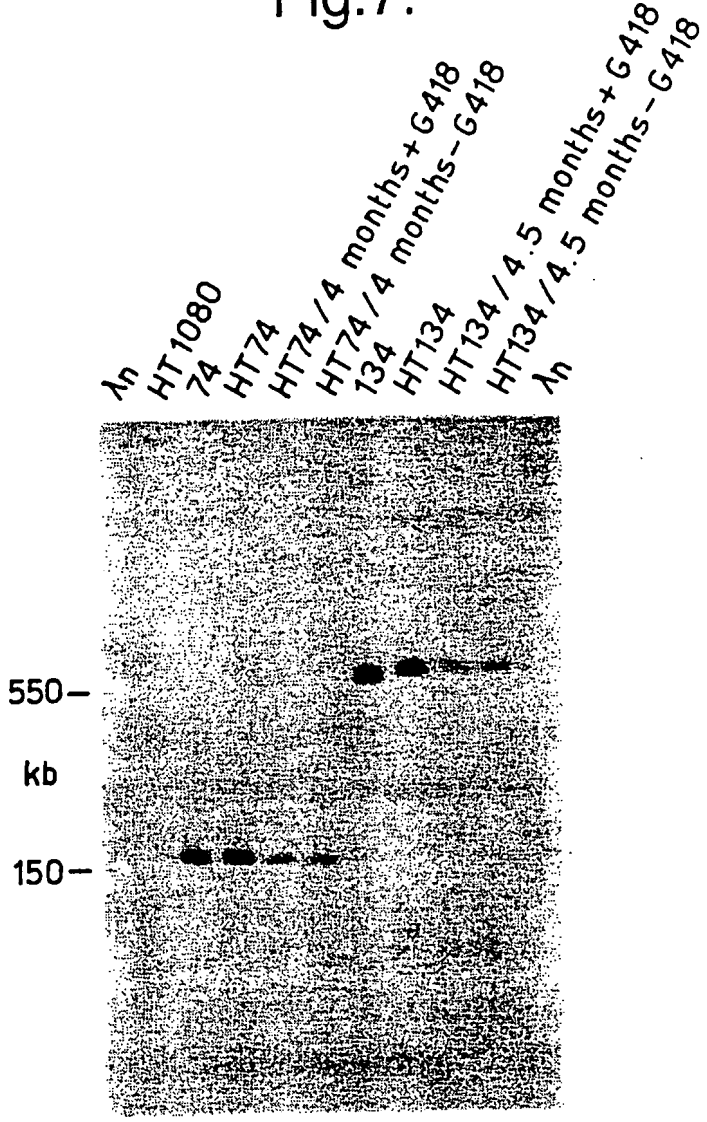

λn
HT1080
74
HT74
HT74/4 months+G418
HT74/4 months-G418
134
HT134
HT134/4.5 months+G418
HT134/4.5 months-G418
λn
550—
kb
150—

Fig.8.

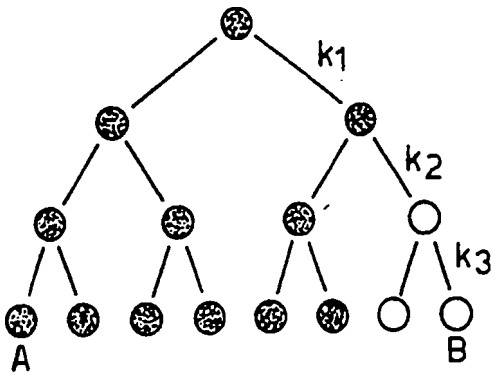

k1
k2
k3
A
B

Fig.9a.

Construction of a prototype MAC vector; Δ1

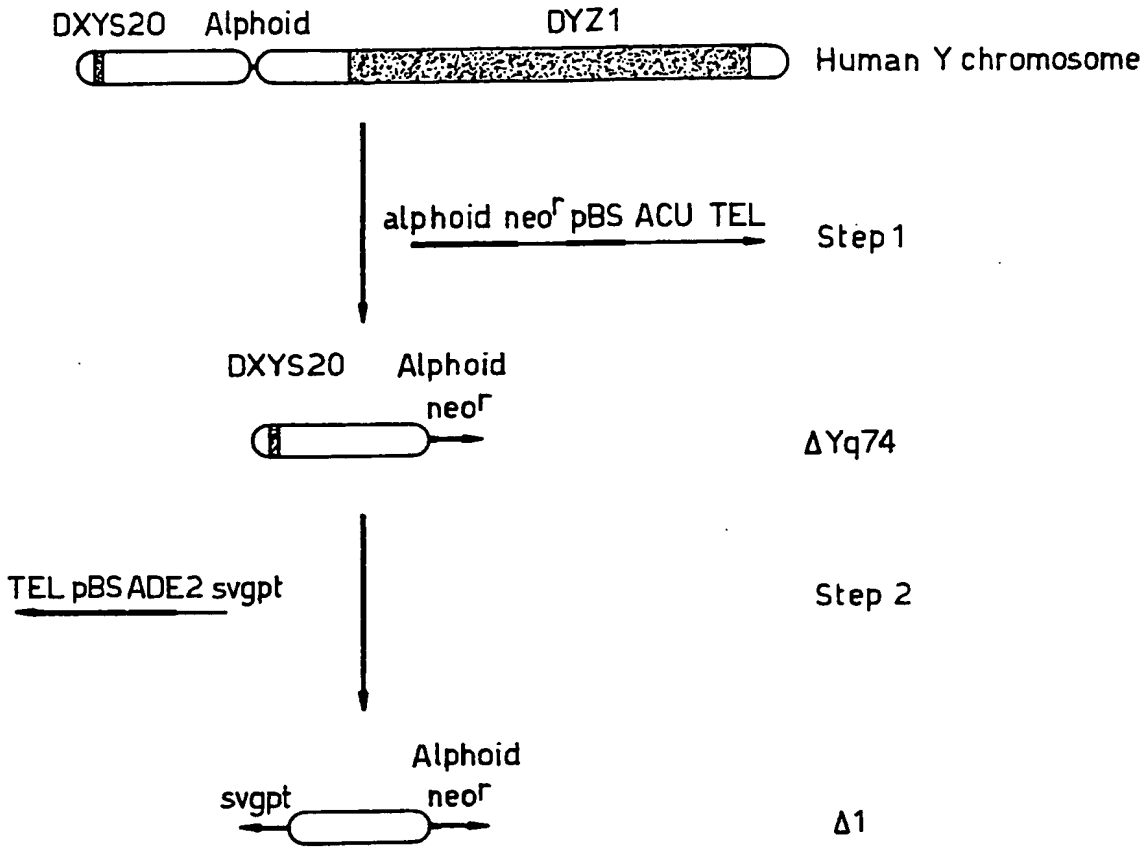
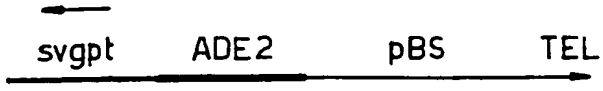


Fig.9b.



1kb

Fig.10.

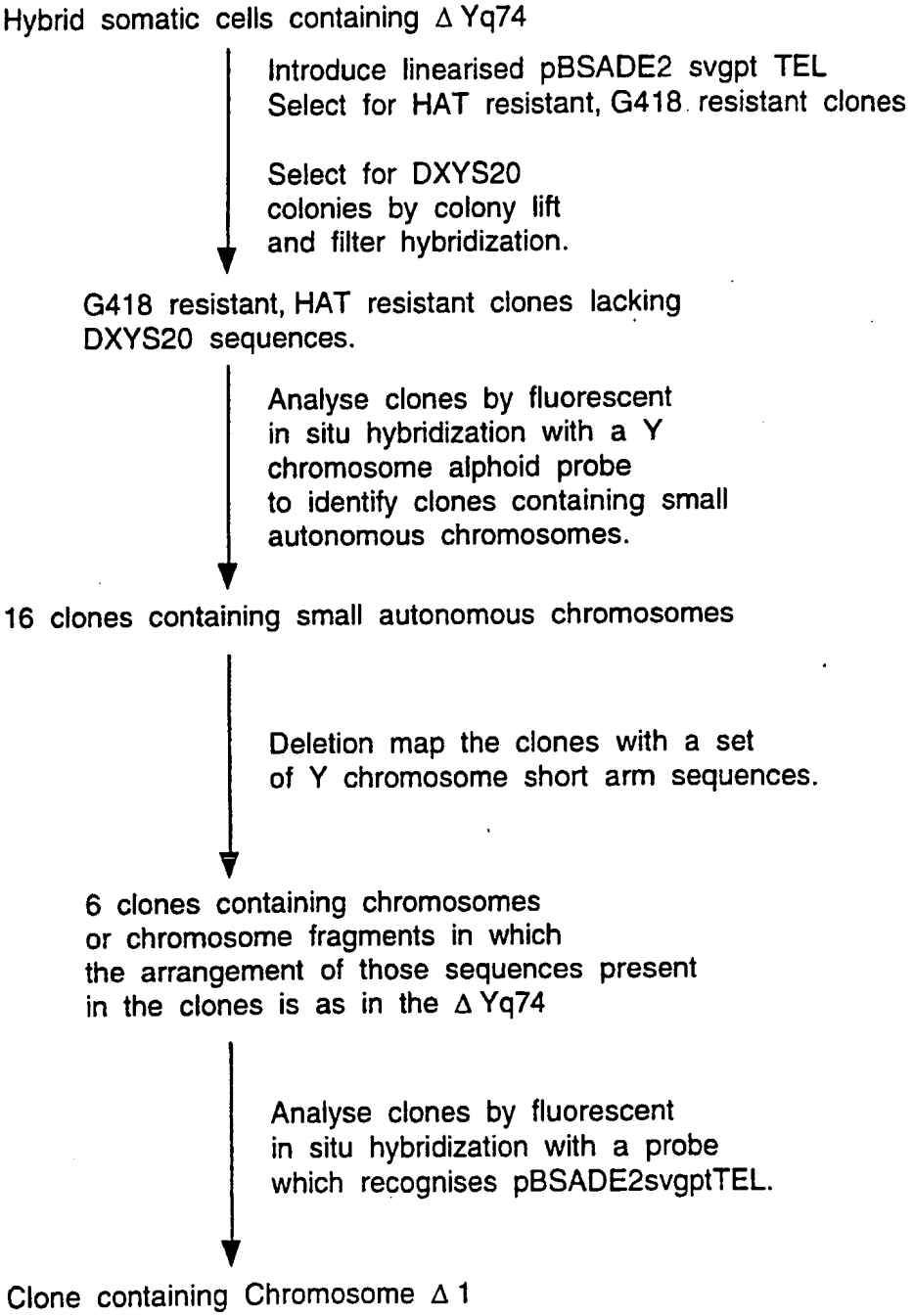

Hybrid somatic cells containing Δ Yq74
Introduce linearised pBSADE2 svgpt TEL
Select for HAT resistant, G418 resistant clones
Select for DXYS20
colonies by colony lift
and filter hybridization.
G418 resistant, HAT resistant clones lacking
DXYS20 sequences.
Analyse clones by fluorescent
in situ hybridization with a Y
chromosome alphoid probe
to identify clones containing small
autonomous chromosomes.
16 clones containing small autonomous chromosomes
Deletion map the clones with a set
of Y chromosome short arm sequences.
6 clones containing chromosomes
or chromosome fragments in which
the arrangement of those sequences present
in the clones is as in the Δ Yq74
Analyse clones by fluorescent
in situ hybridization with a probe
which recognises pBSADE2svgptTEL.
Clone containing Chromosome Δ 1

Fig.12.

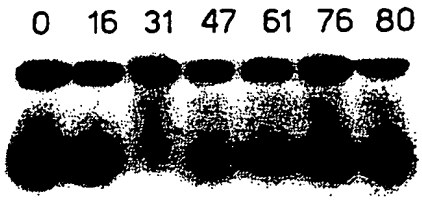
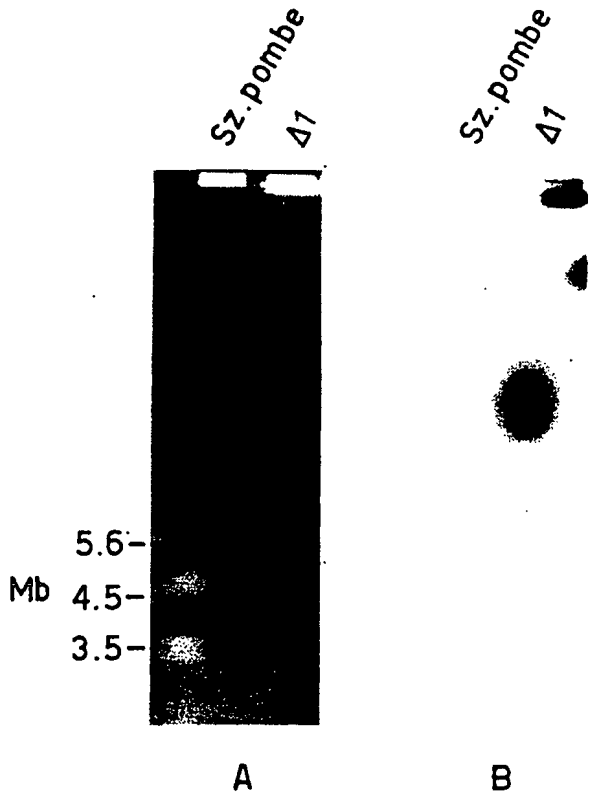


Fig.13.

Fig.14.



INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 95/01195

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/85 C12N15/63 C12N15/64 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CURR. OPIN. GENET. DEV., vol. 2, 1992 pages 479-486, BROWN,W.R.A. cited in the application see the whole document ---	1-20
X	NUCLEIC ACIDS RESEARCH, vol. 21, 1993 pages 27-36, BARNETT,M.A. ET AL. cited in the application see the whole document ---	1-20
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 September 1995	23.10.95
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Olsen, L

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NATURE GENET., vol. 5, 1993 pages 368-375, TYLER-SMITH, C. ET AL. cited in the application see page 368 see page 373 - page 375	1-8, 14-20
X	PROC. NATL. ACAD. SCI. U.S.A., vol. 82, 1985 pages 2106-2110, SUROSKY, R.T. AND TYE, B.-K. cited in the application see page 2106 see page 2109 - page 2110	1,6-9
X	US,A,4 889 806 (OLSON ET AL.) 26 December 1989 see the whole document	1,6-9
X	WO,A,89 09219 (THE GENERAL HOSPITAL CORP.) 5 October 1989 see page 7 - page 9; claims 1-35; figures 7,10	1,6,7, 14,15,19
X	PROC. NATL. ACAD. SCI. U.S.A., vol. 87, 1990 pages 1300-1304, PAVAN, W.J. ET AL. see the whole document	1,6,7
A		8-15
X,P	HUM. MOL. GENET., vol. 3, no. 8, August 1994 pages 1383-1386, TAYLOR, S.S. ET AL. see the whole document	1-7
P,X	WO,A,95 03400 (JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE) 2 February 1995 see claims 1-36	1,6,7, 14,15

1

INTERNATIONAL SEARCH REPORT

Internati Application No
PCT/GB 95/01195

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
US-A-4889806	26-12-89	NONE		
WO-A-8909219	05-10-89	AU-B-	622900	30-04-92
		AU-B-	3346489	16-10-89
		EP-A-	0338266	25-10-89
		JP-T-	3503599	15-08-91
		US-A-	5270201	14-12-93
WO-A-9503400	02-02-95	NONE		